# Hidden Markov model analysis reveals better eye movement strategies in face recognition

**Tim Chuk (u3002534@connect.hku.hk)**
Department of Psychology, The University of Hong Kong, Pokfulam Road, Hong Kong
**Antoni B. Chan (abchan@cityu.edu.hk)**
Department of Computer Science, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong
**Janet Hsiao (jhsiao@hku.hk)**
Department of Psychology, The University of Hong Kong, Pokfulam Road, Hong Kong

**Abstract**

Here we explored eye movement strategies that lead to better performance in face recognition with hidden Markov models (HMMs). Participants performed a standard face recognition memory task with eye movements recorded. The durations and locations of the fixations were analyzed using HMMs for both the study and the test phases. Results showed that in the study phase, the participants who looked more often at the eyes and shifted between different regions on the face with long fixation durations had better performances. The test phase analyses revealed that an efficient, short first orienting fixation followed by a more analytic pattern focusing mainly on the eyes led to better performances. These strategies could not be revealed by analysis methods that do not take individual differences in both temporal and spatial dimensions of eye movements into account, demonstrating the power of the HMM approach.

**Keywords:** hidden Markov model; fixation duration; eye movement; face recognition.

## Introduction

In our daily lives, we frequently fixate at objects that attract our attention. Fixations are usually not extremely short, because during the time when we fixate at an object, the brain needs to analyze the fixation target as well as to plan for the next move. Previous studies showed that the former may take about 100 to 150 msec (Erikson & Erikson, 1971), while the latter may take about 150 to 200 msec (Becker & Jurgens, 1979). It is therefore argued that duration could be used to reveal underlying cognitive processes.

It was found in reading experiments that fixation duration is the best predictor for word frequency and word complexity (Rayner, 1998). Similar findings were also discovered in the context of scene perception. A number of studies showed that the more informative objects in scenes received more and longer fixations (see Unema et al., 2005).

However, in the context of face recognition, few studies have extensively studied the influence of fixation duration. A number of studies (e.g., Henderson et al., 2005; Kelly et al., 2011) documented fixation durations, but not many in-depth analyses were carried out. Schwarzeret al. (2005) conducted an experiment in which they required participants to categorize face images into different categories based on either the overall similarities of all the facial features or the similarity of a specific facial feature. The former intended to manipulate participants to adopt a holistic viewing strategy, while the latter intended to manipulate participants to adopt an analytic viewing strategy. Results showed the two strategies yielded different average gaze durations on each facial region. Hsiao & Cottrell (2008) found that when participants were asked to learn and identify faces, the fixation durations in the study phase gradually increased but the fixation durations in the test phase did not, suggesting that the strategies participants adopted during the two phases were somewhat different.

These findings suggest that duration might be informative to understanding face recognition in addition to fixation locations. However, there are barriers to more extensive analyses. First, the strategies used when looking at faces might not be that straightforward. It would be a simplistic assumption that only one strategy is employed under various conditions. In fact, some visual search studies (e.g., Hooge & Erkelens, 1996) showed that the expected difficulty of the task plays a significant role in determining the length of the fixation durations. It is possible that the fixation durations are task-dependent.

On the other hand, it has also been shown that there are a lot of individual differences in people's eye movement patterns. For example, Castelhano et al., (2009) found that in terms of fixation locations when viewing images, within-participant consistencies were higher than between-participant consistencies. In terms of fixation duration, some studies suggested that there are substantial and persistent between-subject differences among infants (Colombo et al., 1995). It was even found that this difference can be used to predict individual's cognitive abilities. Sigman et al., (1991) discovered that infants with higher cognitive abilities tended to have shorter fixations.

In our previous study (Chuk et al., 2014a), we showed that by analyzing fixation locations with hidden Markov models, we could capture individual differences in eye movement patterns. We also showed that by clustering the HMMs into subgroups, different eye movement strategies could be discovered without making any presumptions. Our result suggested that people who exhibited analytic fixation patterns performed significantly better than their counterparts who exhibited holistic and semi-holistic fixation patterns (Chuk et al., 2014b). However, the previous study did not make use of duration information in the eye fixation data, nor did we look at the study phase fixations. Here, we make use of the duration and the study phase data to further understand eye movement strategies that lead to better performance in face recognition. We explore the duration information with HMMs and draw a comparison between these results and those from only location analyses. We train

an HMM on each participant using the duration of the fixations and cluster the HMMs into groups based on their similarities. We also train an HMM on each participant using the location of the fixations and cluster the HMMs into groups. Our aim is to discover fixation patterns, in both the study and test blocks, that lead to better recognition performance.

## Method

### Participants

We recruited 48 participants, including 24 Chinese (7 males, mean age 21.5, SD = 2.2) and 24 Caucasians (6 males, mean age 21.2, SD = 7.5). Participants were given course credit or honorariums. Asian participants were students at the University of Hong Kong. Of the Caucasian participants, three were exchange students or staff at HKU; the remainders were students at the University of Western Australia. All participants reported normal or corrected-to-normal vision.

### Materials

The stimuli were a total of 56 colored frontal-view faces (28 Asians and 28 Caucasians). The Chinese faces were collected by Professor William Hayward at the University of Hong Kong and the Caucasian faces were collected by Professor Elinor McKone at Australian National University. Half of the faces were females and half of them were males. All the faces were with neutral expressions and were unfamiliar to the participants. The faces were cropped around the chin and ears, and some hair remained visible. The vertical height of the faces was 384 pixels. The horizontal widths varied due to natural variations in face shape and were on average 298 pixels.

During the experiment, the faces were shown on a 22'' CRT monitor with a resolution of 1024 x 768 pixels that was located approximately 60 cm from the participants, which therefore allowed each image to subtend about 8 degrees of visual angle horizontally and 13 degrees vertically.

### Procedure

The experiment consisted of two sessions, each with a study block and a test block. Participants were allowed to take a short break between the sessions. In the study block, participants were asked to remember 14 faces of the same race. A total of 28 faces were selected for the study block and split into two sets of 14 faces (7 males and 7 females). The two sets were counterbalanced across participants. In the test block that followed, participants were asked to recognize the 14 target faces among 14 foils.

Half of the participants viewed the Caucasian faces in the first session and Asian faces in the second session. The other half viewed the Asian faces in the first session and Caucasian faces in the second session.

We recorded participants' eye movements with EyeLink 1000 eye tracker. We used the default settings of the system and performed a nine-point calibration procedure at the beginning of each block. Participants put their heads on a chin rest in order to stabilize their head movements.

In the study blocks, each trial began with a fixation circle at the center of the screen for drift correction. Participants were told to fixate at the circle to ensure that they were looking at the center of the screen when the faces appear. Trials were initiated by the experimenter once the participant was fixating on the circle. A target face was then presented at one of the four quarters of the screen for 5 seconds. In the test blocks, the flow of the trials were the same as that in the study blocks except that the images remained on the screen until the participants responded by pressing one of the two keys that indicated whether they remembered the face or not. No feedback was given during the experiment.

### Hidden Markov model

Hidden Markov models (HMMs) are time-series models that assume dependency between the current state of a time-series data and its previous state. The states of the data are hidden and can be estimated from the probabilistic association between the observed data and the states as well as from the transitional probabilities between the states. An HMM contains a vector of prior values, which indicates the probability of the time-series beginning with each particular state, a transition matrix, which specifies the transition probabilities between any two hidden states, and a Gaussian emission for each state, which captures the probabilistic association between the observed data (e.g. eye fixation duration) and a hidden state.

We trained one HMM per participant using the fixation durations collected from all the trials in either the study or the test blocks. Hence, each participant had two HMMs, one trained on the study block trials and one the test block trials. The hidden states represent clusters of durations that are of similar length. We then categorized the HMMs trained on the same blocks into groups using the variational hierarchical EM algorithm (VHEM) (Coviello et al., 2014). The VHEM algorithm categorizes the input HMMs into groups based on their similarities. It also produces a representation HMM that summarizes the commonalities of the HMMs within the group. The procedure was performed on the study HMMs and test HMMs separately. For both the study and the test blocks, we compared the recognition performance of the groups discovered by the VHEM algorithm. Performance in the study blocks was measured by checking whether the faces appeared in the study trials were identified correctly in the test trials. Performance in the test blocks was measured with d-prime.

For each group with similar fixation durations, we also trained one HMM per participant and per block using the fixation locations, and combined the individual HMMs into a single representative HMM using VHEM to examine the general fixation patterns in the group. We thus visualized how the groups, which differed in their fixation durations, also differed in the spatial distribution of the fixations.

Then, we trained one HMM per individual using only the location information. The location HMMs were also catego-

rized into groups using the VHEM algorithm. We compared the performance of the groups found by categorizing the location HMMs to see whether, with location information alone, the groups also show difference in performance.

## Results

### Categorization of the individual duration HMMs

We clustered the 48 study block HMMs into three groups using the VHEM algorithm. The left side of Figure 1 shows the representation HMM of each group. Table 1 shows that there were approximately same numbers of Caucasians and Asians in each group. No difference in racial distribution over the groups was observed ( $\chi^2(2) = 0.12$, $p = .94$).

It can be seen that the three groups had different patterns of fixation durations. The first group usually began with a relatively short fixation, which was then followed by either another almost equally short or a much longer fixation. The second group showed a different pattern. The first fixations were also short. However, the following fixations were likely to be slightly longer, with 1/3 of the chance being much longer. The third group was distinctive in that there was no long fixation; all the three clusters were within the range of 0 to 500 msec. We found that the first fixations of Group 2 (210.1) were significantly longer than those of Group 3 (153.0), $t(40) = 3.75$, $p < .01$. The difference between Group 1 (180.1) and Group 2 (210.1), $t(15) = 0.99$, $p = .33$, and the difference between Group 1 (180.1) and Group 3 (153.0), $t(35) = 1.49$, $p = .14$, were not significant.

We compared the performance of the groups and discovered that Group 2 performed significantly better than the other two groups. The hit rate of Group 2 ($M = .85$) was significantly better ($F(2, 45) = 4.9$, $p = .01$) than Group 1 ($M = .68$, $t(15) = 2.98$, $p = .01$), and Group 3 ($M = .78$, $t(40) = 2.22$, $p = .03$).

Similarly, we also clustered the 48 test block HMMs into three groups. The right side images of Figure 1 show the representation HMM of each group. Again, there were roughly same numbers of Caucasians and Asians in each group ( $\chi^2(2) = 0.16$, $p = .92$; see Table 2).

The differences between the three groups were less obvious in the test blocks. The three duration clusters of the Group 1 were highly overlapped. Groups 2 and 3 had rather clear distinction between one and the other two clusters. As can be seen from the transition matrices, all groups began with short fixations, which are followed by some relatively longer fixations. However, the durations in clusters of Group 3 were shorter than that of Group 2. We found that the first fixations of Group 3 (129.9) was significantly shorter than that of Group 1 (189.2), $t(35) = 7.62$, $p < .01$, and that of Group 2 (187.3), $t(31) = 4.96$, $p < .01$. Regarding the average duration of all fixations, Group 2 (297.6) was significantly longer than Group 1 (244.9), $t(24) = 6.49$, $p < .01$, and Group 3 (214.9), $t(31) = 9.39$, $p < .01$. Comparison of the recognition performances (d' prime) showed a significant effect of group, $F(2, 45) = 3.59$, $p = .04$; the results suggested that the Group 3 ($M = 2.03$) performed significantly better than Group 1 ($M = 1.49$, $t(35) = 2.69$, $p = .01$), but not Group 2 ($M = 1.85$, $t(31) = -.9$, $p = .37$).

| | Caucasians | Asians |
|---|---|---|
| Group 1 | 3 | 3 |
| Group 2 | 5 | 6 |
| Group 3 | 16 | 15 |

Table 1: distributions of the two races over the three duration groups (study block).

| | Caucasians | Asians |
|---|---|---|
| Group 1 | 7 | 8 |
| Group 2 | 6 | 5 |
| Group 3 | 11 | 11 |

Table 2: distributions of the two races over the three duration groups (test block).

### Analysis of fixation locations for duration groups

Using the groups uncovered above, we then estimated a representative HMM for each group using the fixation locations. Figure 2 shows representation HMMs for the study blocks (left) and the test blocks (right).

The images show that for study blocks, Group 1 usually just fixated at everywhere on a face rather than any specific facial feature. Groups 2 and 3, on the other hand, usually began with fixations that help them to locate the face on the screen, which was then followed by fixations concentrated more on the eyes. The difference between the two groups can be observed from the transitions. Group 2 explored different parts of a face, while Group 3 tended to stare at only some parts of a face. For Group 2, the transition probability from the red cluster to green cluster was about the same as staying in red (~47%), and likewise when starting in the green cluster. For Group 3, the probability to stay in red was almost twice that of a transition from red to green, and similarly for the green cluster.

The test block HMMs show somewhat different results. Although the spatial distributions of the three clusters in each group seemed different, the transition matrices suggest that the differences were subtle. All three groups began with a fixation without any specific target, which was followed by fixations either again without a target facial feature or more concentrated on the eyes. The chances of two options are about 50-50. The result suggests that when participants were required to identify the faces, all different groups paid more attention to the eyes.

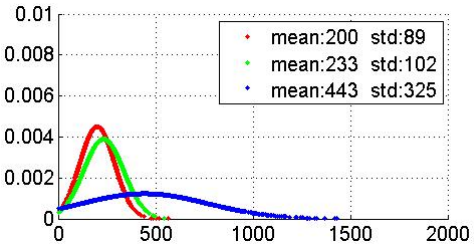
### Categorization of the individual location HMMs

For the study block, we clustered the 48 location HMMs into three groups. The left images of Figure 3 show the study block representation HMMs of each group. Table 3 shows the numbers of Caucasians and Asians in each group. There were significantly more Asians than Caucasians in Group 3 ( $\chi^2(2) = 6.57$, $p = .04$).

Group 1 looked mainly at the whole face without a specific target. There were occasionally fixations to the lower part of the face (transition from red to blue is 10%). Group 2 and Group 3 both spent more time on the eyes. However,

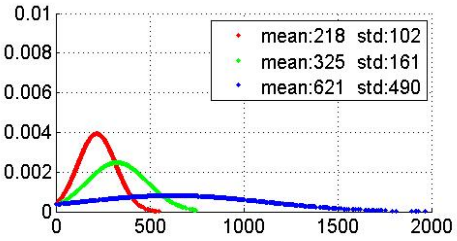
while Group 3 mainly started from some random areas on a face and then consistently transited between the two eyes and other areas, Group 2 in some cases were likely to begin with and to remain staring at the eyes. Comparison of the recognition performances (hit rate) shows a significant effect of group, $F(2, 45) = 3.65$, $p = .03$; the results suggested that Group 3 ($M = .83$) performed significantly better than Group 1 ($M = .73$, $t(29) = 2.54$, $p = .02$), but not Group 2 ($M = .79$,

**Study Block**



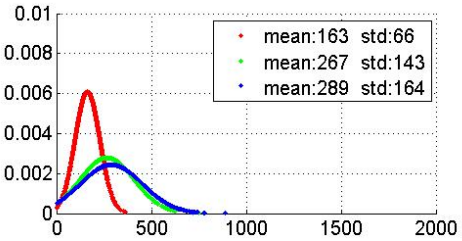

Group 1

| from\to | red cluster | green cluster | blue cluster |
|---|---|---|---|
| priors | .79 | .09 | .12 |
| red cluster | .30 | .35 | .35 |
| green cluster | .22 | .44 | .34 |
| blue cluster | .15 | .55 | .30 |




Group 2

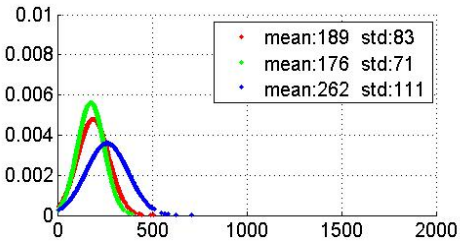
| from\to | red cluster | green cluster | blue cluster |
|---|---|---|---|
| priors | .91 | .02 | .07 |
| red cluster | .02 | .68 | .30 |
| green cluster | .08 | .58 | .34 |
| blue cluster | .08 | .62 | .30 |




Group 3

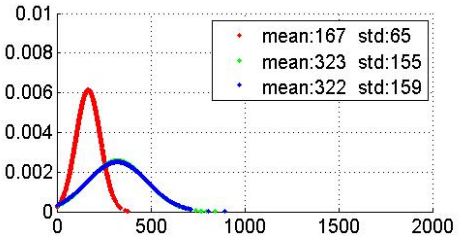
| from\to | red cluster | green cluster | blue cluster |
|---|---|---|---|
| priors | .89 | .03 | .07 |
| red cluster | .03 | .76 | .21 |
| green cluster | .04 | .69 | .27 |
| blue cluster | .04 | .62 | .34 |

**Test Block**



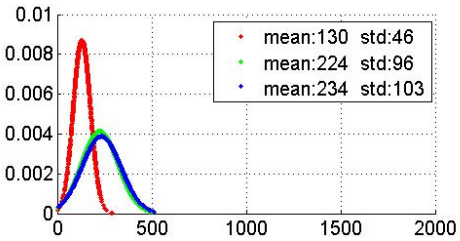

Group 1

| from\to | red cluster | green cluster | blue cluster |
|---|---|---|---|
| priors | .72 | .20 | .08 |
| red cluster | <.01 | <.01 | .99 |
| green cluster | <.01 | <.01 | .99 |
| blue cluster | .04 | .01 | .95 |




Group 2

| from\to | red cluster | green cluster | blue cluster |
|---|---|---|---|
| priors | .78 | .04 | .18 |
| red cluster | <.01 | .16 | .84 |
| green cluster | .02 | .04 | .94 |
| blue cluster | .04 | .01 | .95 |

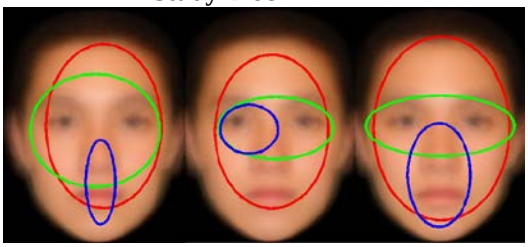



Group 3

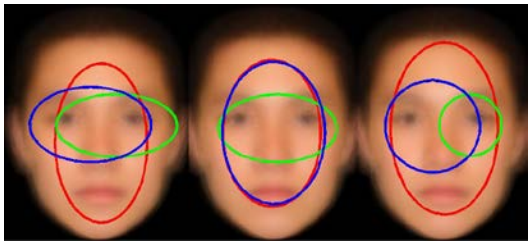
| from\to | red cluster | green cluster | blue cluster |
|---|---|---|---|
| priors | .91 | .04 | .05 |
| red cluster | .01 | .60 | .39 |
| green cluster | .01 | .20 | .79 |
| blue cluster | .01 | .36 | .63 |

Figure 1. The representation HMMs of the three groups based on clustering fixation durations from the study (left) and test blocks (right). The curves indicate the clusters. X-axis represents fixation duration (in msec). Y-axis represents the probability densities. The red cluster has the highest prior probability, i.e., most likely contains the first fixation. The blue cluster has the largest variance. In the study block, Group 2 performed the best. In the test block, Group 2 and 3 outperformed Group 1.

**Study Block**



**Test Block**

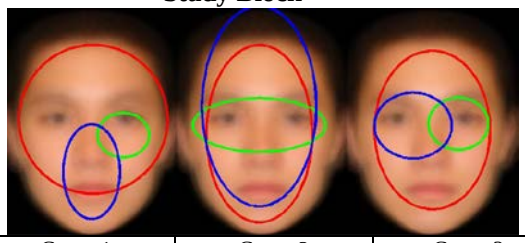


| | Group 1 | | | Group 2 | | | Group 3 | | | | Group 1 | | | Group 2 | | | Group 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| from\to | r | g | b | r | g | b | r | g | b | from\to | r | g | b | r | g | b | r | g | b |
| p | .97 | .02 | .01 | .96 | .03 | .01 | .82 | .14 | .04 | p | .73 | .03 | .24 | .55 | .01 | .44 | .87 | .06 | .07 |
| r | .59 | .38 | .03 | .47 | .46 | .07 | .62 | .35 | .03 | r | .51 | .25 | .24 | .12 | .54 | .34 | .54 | .24 | .22 |
| g | .40 | .50 | .10 | .56 | .42 | .02 | .34 | .58 | .08 | g | .57 | .07 | .36 | .13 | .40 | .47 | .38 | .13 | .49 |
| b | .61 | .07 | .32 | .30 | .55 | .15 | .37 | .27 | .36 | b | .51 | .33 | .16 | .34 | .47 | .19 | .45 | .41 | .14 |

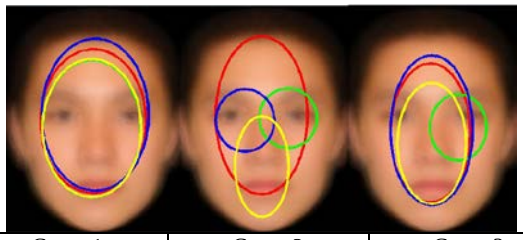
** p = prior; r = red; g = green; b = blue*

Figure 2. The representation HMMs for fixation locations of the three duration groups for study (left) and test blocks (right).

**Study Block**



**Test Block**



| | Group 1 | | | Group 2 | | | Group 3 | | | | Group 1 | | | | Group 2 | | | | Group 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| from\to | r | g | b | r | g | b | r | g | b | from\to | r | g | b | y | r | g | b | y | r | g | b | y |
| p | .95 | .01 | .04 | .68 | .21 | .11 | .93 | .05 | .02 | p | .45 | .11 | .35 | .09 | .95 | .03 | .01 | .01 | .41 | .18 | .32 | .09 |
| r | .89 | .01 | .10 | .44 | .51 | .05 | .47 | .39 | .14 | r | .33 | .60 | .01 | .06 | .15 | .35 | .43 | .07 | .09 | .60 | .04 | .27 |
| g | .40 | .25 | .35 | .36 | .60 | .04 | .46 | .18 | .36 | g | .11 | .40 | .05 | .44 | .08 | .12 | .58 | .22 | .30 | .23 | .14 | .33 |
| b | .55 | .09 | .36 | .33 | .55 | .12 | .28 | .55 | .17 | b | .18 | .23 | .09 | .50 | .09 | .64 | .09 | .18 | .10 | .60 | .05 | .25 |
| | | | | | | | | | | y | .23 | .42 | .05 | .30 | .19 | .22 | .41 | .18 | .10 | .51 | .12 | .27 |

** p = prior; r = red; g = green; b = blue; y = yellow*

Figure 3. The representation HMMs of the 3 groups clustered using fixation locations from study (left) and test blocks (right, from Chuk et al., 2014b). In study, Group 2 and 3 outperformed Group 1. In test, Group 2 performed the best.

t(29) = 1.09, p = .28). Difference between Groups 1 and 2 was marginal (t(32) = 1.72, p = .09). For the test block, we performed similar analysis in our previous study (Chuk et al., 2014b) and results are presented in Figure 3 (right). We trained one HMM per participant and categorized them into three groups using the VHEM algorithm. We found that there were roughly equal numbers of Caucasians and Asians in Groups 1 and 2, but more Asians than Caucasians in Group 3 ( $\chi^2(2) = 4.55$, p = .03).

| | Caucasians | Asians |
|---|---|---|
| Group 1 | 10 | 7 |
| Group 2 | 11 | 6 |
| Group 3 | 3 | 11 |

Table 3: distributions of the two races over the three groups clustered using fixation location (study block).

| | Caucasians | Asians |
|---|---|---|
| Group 1 | 9 | 5 |
| Group 2 | 13 | 11 |
| Group 3 | 2 | 8 |

Table 4: distributions of the two races over the three groups clustered using fixation locations(test block).

Group 1 demonstrated a holistic viewing pattern; no facial feature was looked at in specific. Group 2 looked at the two eyes respectively after the first fixations, and in some cases paid some attention to the mouth (~7%). Group 3 showed a similar pattern to that of Group 1, but with more attention devoted only to the right eye. Comparison of the recognition performances (d' prime) discovered that a significant effect of grouping, F(2, 45) = 5.86, p = .01, results suggest that Group 2 (M = 2.1) performed significantly better than Group 1 (M = 1.53, t(36) = 2.81, p = .01), and Group 3 (M = 1.54, t(32) = 2.89, p = .01).

## Discussion

The results have shown that duration is an informative source of information about the strategies participants adopted that lead to better performance in face recognition, in addition to fixation location information.

By grouping participants into three groups using only the duration information, we have discovered a significant difference in the performance of the participants. In the study blocks, the hit rate was the best for Group 2, which had almost exclusively short fixations at beginning and longer fixations afterwards. The spatial distribution of the fixations shows that Group 2 was more likely than Group 3 to shuffle between ROIs, indicating that best performance was achieved by looking at different parts of a face (with long fixations) than staring at only one facial feature. Grouping of the location HMMs have shown a similar tendency. People who looked at other areas as well as focused on the eyes performed better than the others.

The result that all the participants showed shorter first fixations and longer following fixations was also consistent with that reported by Hsiao & Cottrell (2008), in which it was found that for the study phase, the first fixations were shorter and the following ones gradually increased in duration. With the HMMs, we found that although the pattern was shared by all the participants, differences in duration and transition patterns among the subgroups were associated with performance difference.

For the test blocks, although all 3 groups of the duration HMMs generally had short first fixations, comparison between groups showed that for Group 3, which had the best performance, their first fixations were significantly shorter than those in Group 1 and 2. The location HMMs shown in

Figure 3 (right) have revealed that those who looked specifically at the two eyes (Group 2 in Figure 3) performed better in the test blocks. The performance of those who looked only at the right eye (Group 3) was the same as the holistic viewing people (Group 1). Since the first fixation is usually for locating the face, together with the duration analysis, this may imply that some of the better performing people have a more consistent strategy to locate faces on the screen, so that their first fixations are shorter, and then subsequently look at both eyes for face identification. Yet, some participants (Group 2 of the duration HMMs in Figure 1) had longer average fixation duration than the rest but achieved level of performance comparable to that of Group 3, suggesting the existence of individual differences.

The observation that in the test block most participants spent shorter time on fixating at the faces than the study block is contradictory to the finding from the study blocks, which showed that people whose fixations were all short tended to perform worse (Group 3 of duration HMMs in the study blocks in Figure 1). This may indicate that different strategies had been adopted for the study and test blocks. It is convergent to Hsiao and Cottrell (2008) and Henderson et al., (2005), which showed that the strategies participants adopted in different tasks were different.

The current study also demonstrates the advantages of using HMMs to analyze fixation durations. HMMs contain clusters and their transition information. The clusters categorize durations of similar length into same clusters, and the transition information reveals the transitions between clusters, so that the duration information in eye movement data could be more precisely interpreted. Moreover, HMMs are able to capture individual differences because each participant is modeled with an HMM. By clustering the HMMs into groups, similar viewing strategies shared by a number of participants could be discovered from data. For example, the test block categorization result shows that good performance could be achieved by two different strategies, one with shorter and one with longer fixation durations. The discovery of the long fixation strategy is only possible when individual differences have been taken into account during the analyses.

In summary, the findings suggest that duration can be an informative source to understanding eye movements in face recognition. It distinguishes good performers from bad performers in both study and test blocks. Together with the analyses with location information, our results suggest that an exploratory strategy looking at different facial features (especially the eyes) with long fixations is beneficial in the study phase, whereas in the test phase, an efficient, short first orienting fixation followed by a more analytic pattern focusing mainly on the eyes leads to better performances.

## Acknowledgments

We thank William Hayward and Kate Crookes for sharing the data and details of the experiment design. We are grateful to the Research Grant Council of Hong Kong (project 17402814 to J.H. Hsiao and CityU 110513 to A.B. Chan).

## References

Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module?Evidence from human neonates. *Infant Behavior and Development*, *23*(2), 223-229.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision research*, *19*(9), 967-983.

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*(3), 6.

Colombo, J., Freeseman, L. J., Coldren, J. T., & Frick, J. E. (1995). Individual differences in infant fixation duration: Dominance of global versus local stimulus properties. *Cognitive Development*, *10*(2), 271-285.

Coviello, E., Chan, A. B., & Lanckriet, G. R. (2014). Clustering hidden Markov models with variational HEM. *The Journal of Machine Learning Research*, *15*(1), 697-747.

Chuk, T., Chan, A. B., & Hsiao, J. H. (2014a). Understanding eye movements in face recognition using hidden Markov models. *Journal of vision*, *14*(11), 8.

Chuk, T., Luo, A.X., Crookes, K., Hayward, W.G., Chan, A. B., & Hsiao, J. H. (2014b). Caucasian and Asian eye movement patterns in face recognition: A computational exploration using hidden Markov models. *Journal of vision*, *14*(10), 1212.

Eriksen, C. W., & Eriksen, B. A. (1971). Visual perceptual processing rates and backward and forward masking. *Journal of Experimental Psychology*, *89*(2), 306.

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & cognition*, *33*(1), 98-106.

Hooge, I. T. C., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics*, *58*(7), 969-976.

Hsiao, J. H. W., & Cottrell, G. (2008). Two fixations suffice in face recognition.*Psychological Science*, *19*(10), 998-1006.

Kelly, D. J., Liu, S., Rodger, H., Miellet, S., Ge, L., & Caldara, R. (2011). Developing cultural differences in face processing. *Developmental science*,*14*(5), 1176-1184.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, *124*(3), 372.

Schwarzer, G., Huber, S., & Dümmler, T. (2005). Gaze behavior in analytical and holistic face processing. *Memory & Cognition*, *33*(2), 344-354.

Sigman, M., Cohen, S. E., Beckwith, L., Asarnow, R., & Parmelee, A. H. (1991). Continuity in cognitive abilities from infancy to 12 years of age. *Cognitive Development*, *6*(1), 47-57.

Unema, P. J., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, *12*(3), 473-494.